Module 2 Overview

Introduction

> "*Play til your fingers bleed.*"
> *--Bryan Adams, Summer '69.*

The purpose of citing tihs quote is twofold. First, I have to say that machine learning can be seen as a process of learning by repetition. We present data to the machine so it can somehow capture the important features of it. Second, machine learning is not an obscure science reserved to a few gurus, but something that is easy to understand (except a few parts of it, that are not well understood yet) and that anyone can learn just by reading, practicing and being stubborn.

In our first contact with Machine Learning, we are to get familiar with the basic concepts of Statistical Learning Theory. This theoretical framework introduced by V. Vapnik and V. Chervoneniks, and and co-workers, represent the root criteria to construct the very popular Support Vector Machines. But beyond that, this theory is a point of view of Machine Learning wort to understand because of the intuition that this provides about the concepts of learning. The concepts of Empirical Risk and Structural risk are used to construct the SVM, but they give a broad, intuitive view of what goes on with any learning machine in general, so it is important to interiorize these concepts as a general matter.

Bryan Adams (Flicr)

Canadian musician.

Module 2 Summary

- Slides
- Weekly Discussion
- Weekly Quizzes

If you have questions about some aspect of Learn, **UNM LEARN Support** is available to troubleshoot technical problems.

Contact them 24/7 at 505-277-0857, 1-877-688-8817 or use the "Create a Support Ticket" link on the left Course Menu.

Learning Objectives

By completing the lessons, group discussion, homework and quizzes for this module, students will be able to:

- Analyze and interpret the very basics of statistic learning theory (SLT).
- Extend the SLT to the concepts of the Vapnik-Chervonenkis (VC) dimension and interpret and apply the VC dimension to particular cases.
- Reproduce the concepts of empirical risk and structural risk.

Required Instructional Materials

- This module is structured in several instructional sets of slides, plus some supplementary original documents that introduce to the support vector machines (see below), that will be useful in the next modules too.

Activities

In order to assess this module, only three quizzes are required, that will help to self assess the expected achievements for it.

Supplementary materials

- **Support vector machines for classification**: Students are required to download the paper "A Tutorial on Support Vector Machines for Pattern Recognition", by Christopher Burges, prior to the session, and study only sections 1 and 2. Further study will be required later. Alternatively, students can use the book "The Nature of Statistical Learning Theory", by V. Vapnik, Springer, 2010, pp 17 to 55, for a more detailed reading.